arXiv:2601.00142v1 [cs.AI] 1 Jan 2026

AnAI MonkeyGets Grapes for Sure – 球面神经网络实现可靠决策

Tiansi Dong
计算机实验室 b
剑桥大学，英国剑桥 JJ Thomson Ave 15 号 td540@cam.ac.uk

Henry He
菲利普斯 Academy A
安多弗，缅因州安多弗市主街180号，美国
hhe26@andover.edu

Pietro Li`o
剑桥大学计算机实验室，英国剑桥市 JJ Thomson 大道15号
pl219@cam.ac.uk

Mateja Jamnik
计算机 Lab
实验室 剑桥大学，英国剑桥市 JJ Thomson 大道15号
mj201@cam.ac.uk

摘要

本文比较了神经推理的三类方法：大语言模型推理、基于监督学习的推理和基于显式模型的推理。大语言模型仍不可靠，难以完成简单的决策任务，而动物无需经过海量语料训练便能掌握这些任务。通过析取三段论推理测试，我们表明，基于监督学习的推理不如通过显式模型构建进行推理。具体而言，我们表明，经过训练、在经典三段论推理中达到 100.00% 的欧拉网络，也可以训练到在析取三段论推理中达到 100%准确率。然而，重新训练后的欧拉网络遭受了严重的灾难性遗忘（在已学会的经典三段论推理任务中，性能降至 6.25%），其推理能力也仅限于模式层面。我们提出了一个新版本的球面神经网络，将概念表示为n维球体表面上的圆。这些球面神经网络通过补圆表示否定算子，并通过过滤掉会形成不可满足圆形构型的不合逻辑的陈述，实现可靠的决策。我们证明，球面神经网络能够掌握 16 项三段论推理任务，包括严格的析取三段论推理，同时保留经典三段论推理的严谨性。我们得出结论，在神经推理的三类方法中，基于显式模型构建的神经推理最为可靠。

1 引言

可靠决策对于高风险应用至关重要。尽管大语言模型在许多方面都取得了前所未有的成功，例如实现类人交流（Biever，2023年）、下围棋（Silver 等，2017年；Schrittwieser 等，2020年）、预测复杂蛋白质结构（Abramson 等，2024年）或天气预报（Soliman，2024年），但它们在简单推理中仍会出错（Mitchell，2023年）。此外，大语言模型容易出现预测准确、解释却不正确的情况（Creswell 等，2022年；Zelikman 等，2022年；Park 等，2024年），而且尚未达到高风险应用所需的可靠性，即在生物医学中进行决策（Wysocka 等，2025年）。

预印本。正在审稿中。

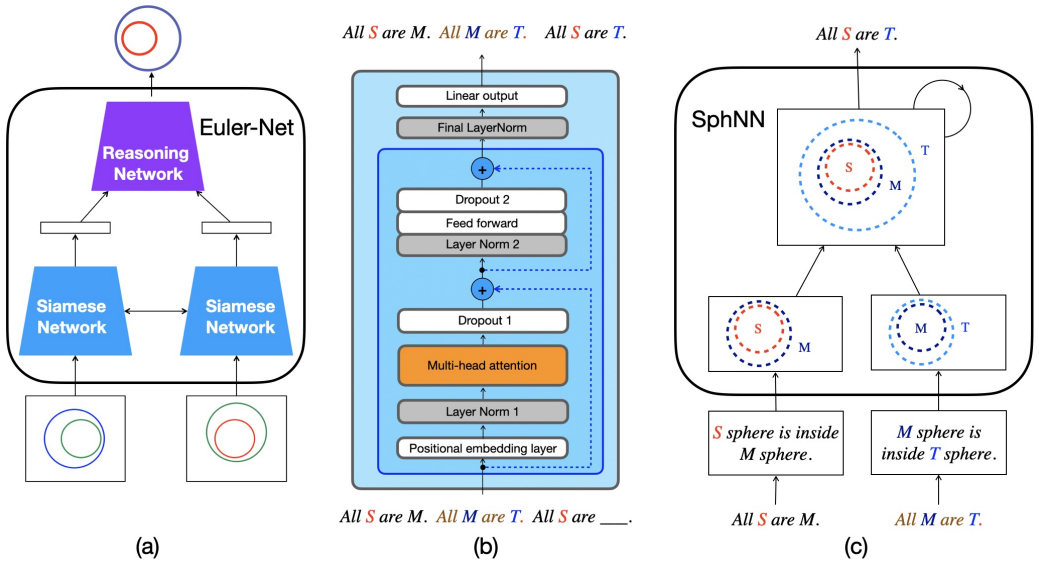

图1：（a）欧拉网络的架构。输入为两张图像，分别表示亚里士多德三段论推理的两个前提。蓝色圆内的绿色圆表示“所有绿色的都是蓝色的”；绿色圆内的红色圆表示“所有红色的都是绿色的”。其输出表示“所有红色的都是蓝色的”。（b）大语言模型的 Transformer 组件学习预测缺失的词语。（c）用于三段论推理的球面神经网络（SphNN）架构。SphNN 将三段论陈述转换为球体之间的空间陈述，构建统一的配置，并通过检查构建出的配置得出结论。

在某些情况下，三段论推理看起来可能简单得令人误以为真，例如以下情况 ：

止血疾病通路的所有成员也都是一般疾病通路的成员。基因 GP1BB 是止血疾病通路的成员。基因 GP1BB 是疾病通路的成员。∴

然而，Eisape等人（2024）表明，尽管大语言模型在三段论推理方面的表现优于普通人类，但其准确率仍仅约为75%，而且较大的模型并不总是优于较小的模型。Lampinen等人（2024）也得出了相近的结论：在三段论等抽象推理任务中，大语言模型在熟悉的情境下可能取得高于随机水平的表现，但在不太熟悉的情境下则存在诸多不足。Wysocka等人（2025）在高风险生物医学背景下，测试了大语言模型在若干种广义三段论推理任务中的表现，例如广义假言推理、析取三段论，结果发现，零样本大语言模型在广义假言推理上的平均准确率为70%，而在析取三段论上仅为23%。至关重要的是，零样本和少样本大语言模型都表现出对表层词汇变化的明显敏感性。

尽管存在这些局限，在三段论任务（及其他任务）上评估大语言模型的推理表现，可以帮助我们洞察（人类）理性的起源（Lampinen 等人，2024年），并有助于寻找实现可靠神经推理与决策的替代方法。大语言模型通过在大规模语言语料上训练，习得了类人交流和推理能力。然而，沟通能力强并不一定意味着推理能力强（Fedorenko 等人，2024年）。相反，推理和决策能力不一定依赖于广泛的语言习得。例如，聪明的猴子能够通过析取三段论决策获得葡萄（Ferrigno 等人，2021年），这表明三段论推理可以通过视觉空间输入来诱发，而不依赖语言能力。尽管近期评估显示，大语言模型仍无法实现稳健的三段论推理，但在本研究中，我们重新审视 Ferrigno 等人（2021年）的实验，以探索替代性的神经网络架构和神经推理方法。我们将用于三段论推理的神经网络分为以下三类（如图 1 所示）：

1. 带图像输入的监督网络。受深层卷积神经网络（CNN）（He 等人，2016年）与视觉皮层之间结构相似性的启发（Yamins 和 Di-Carlo，2016年），欧拉网络是一种特殊的卷积神经网络，能够利用欧拉图风格的图像输入进行亚里士多德三段论推理（Wang等人，2018年、2020年）。

2. 基于语言输入的监督网络。流行的神经网络架构是 Transformer（Vaswani 等人，2017），它是大语言模型的基础（Google，2023年；Touvron 等人，2023年；OpenAI，2023年；Jiang 等人，2023年；Mistral，2023年）。基本的训练方法是使用 Transformer 进行掩码词预测（Raschka，2024年）。给定文本 所有希腊人都是人类。所有人类都是必死的。因此，所有希腊人都是 ，大语言模型经过训练后会预测必死的。

3. 通过显式模型构建进行推理的神经网络。充分的实证研究支持用于推理的模型理论，该理论认为，推理是构建并检查心智模型的过程（Johnson-Laird 和 Byrne，1991；Knauff 等人，2003；Goodwin 和 Johnson-Laird，2005；Knauff，2009、2013）。根据心智模型理论，球面神经网络通过在欧几里得空间或双曲空间中构建欧拉图来进行亚里士多德三段论推理（Dong 等人，2024年、2025年）。事实证明，它们能够达到符号层级逻辑推理的严谨程度。

本文比较了类别1接收图像输入的监督式神经网络与类别3通过显式模型构建实现的神经网络的推理表现（附录中，我们报告了 OpenAI GPT-5 和 GPT-5-nano 在三段论推理方面的表现）。我们以各种类型的三段论式决策对它们进行挑战。这类决策是许多高风险应用的核心，例如法律判决（Deng等人，2023；Constant，2023）和医学诊断（Mushlin 和 Harry L. Greene，2010；Ambags等人，2023）。

本文的贡献有多方面：(1) 通过将欧拉图定义为 n维球体表面上的圆，我们提升了球面神经网络（Dong等人，2024、2025；类别3中的代表性网络）执行16种一阶三段论式推理类型（Betz等人，2021）的表示能力，其中包括析取三段论推理；(2) 我们评估了球面神经网络构建维度为2，3，15，30，100，200，1000，2000，3000，和10000的欧拉图的表现，它在所有这些类型上都达到了100%准确率，表明我们的球面神经网络在所有经典三段论类型中都保持了有效性。(3) 我们改造了欧拉网络（Wang等人，2018、2020；类别1中的代表性网络），使其能够执行析取三段论推理，并达到了100%准确率；(4) 我们创建了新的测试数据，以检验欧拉网络的稳健性，并证明其推理表现受输入模式限制。我们的实验结果表明，类别3网络更适合实现可靠推理与决策。

本文其余部分的结构如下：第2节回顾 Ferrigno 等人（2021年）的猴子实验设计，作为我们实验的背景；第3节综述三段论式推理及相关研究；第4节介绍我们提出的、用于各种三段论推理的球面神经网络版本；第5节对球面神经网络与欧拉网络进行比较研究，并报告实验结果；第6节总结本文，并列出未来可能的研究方向。

2 聪明猴子能够进行析取三段论推理

Ferrigno等人（2021年）对猴子进行了一系列实验，表明聪明的猴子能够进行析取三段论推理。过程简述如下：研究人员在猴子面前放置四个罐子，并用挡板遮住前两个罐子。随后，他们把一颗葡萄放进其中一个罐子里。因此，猴子知道这两个罐子中有一个装着葡萄，但不知道具体是哪一个。接着，他们对最后两个罐子重复这一过程。因此，在猴子看来，每个罐子里有葡萄的概率都是50%。之后，研究人员打开第一个罐子；如果里面有葡萄，就把葡萄给猴子；如果没有，则允许猴子掀开剩下三个罐子中的一个来取葡萄。在这种情况下，聪明的猴子会掀开第二个罐子，因为其中有葡萄的概率为100%，如图2所示。我们可以将这一决策过程抽象为一种析取三段论推理，如下所示。

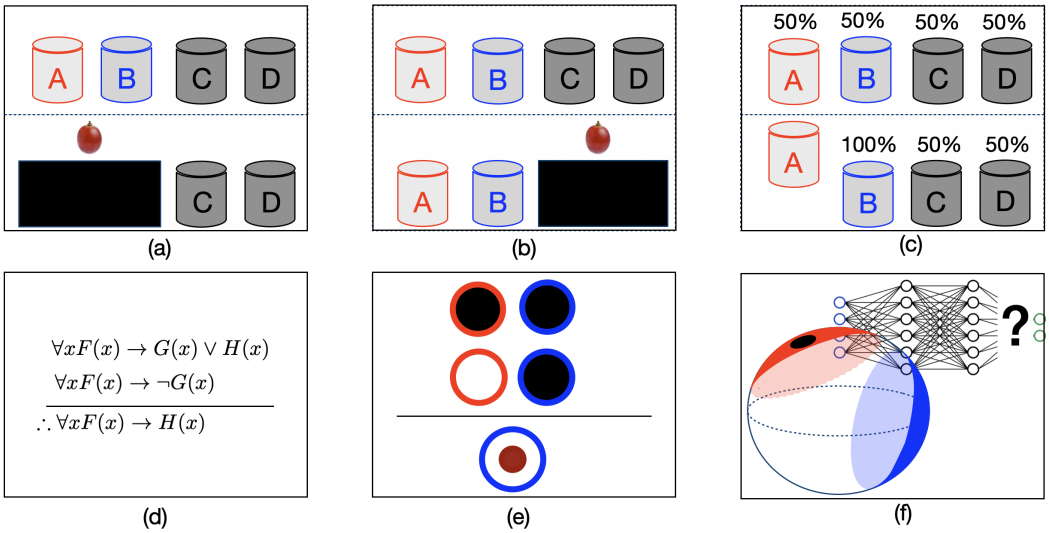

图2：猴子面前有四个罐子，分别是A、B、C和D。(a) 一块黑板挡住罐子A和罐子B；有人将一颗葡萄放入黑板后面的罐子A或罐子B中；(b) 另一块黑板挡住罐子C和罐子D；有人将另一颗葡萄放入黑板后面的罐子C或罐子D中；(c) 移开黑板后，猴子会看到四个罐子，每个罐子里有葡萄的概率都是50%。如果罐子A是空的，聪明的猴子就会抬起罐子B。(d) 聪明猴子的决策过程可以抽象为析取三段论推理；(e) 析取三段论推理的图示表示；(f) 球面神经网络和监督式神经网络可以模拟析取三段论推理。

$\forall x$ 葡萄$(x) \rightarrow$ 在罐子A中$(x) \vee$在罐子B中(x)。
$\forall x$葡萄$(x) \rightarrow \neg$在罐子A中(x)。
$\forall x$葡萄$(x) \rightarrow$ 在罐子B中(x)。 $\therefore$

广义化的版本：

$\forall xG(x) \rightarrow A(x) \vee B(x)$】【。
$\forall xG(x) \rightarrow \neg A(x)$。
$\forall xG(x) \rightarrow B(x)$。 $\therefore$

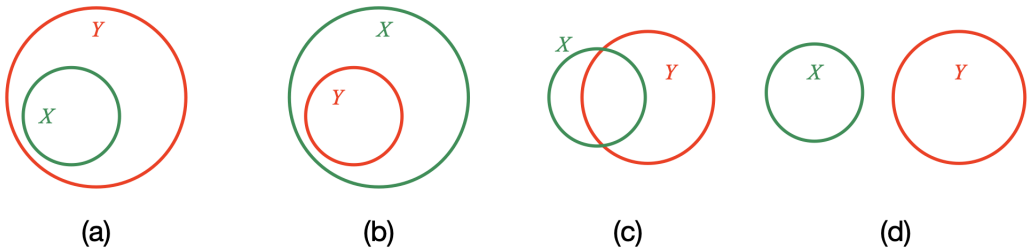

图3：三段论陈述的基本图示表示。(1) 所有 X 都是 Y 表示为 (a) $X \subset Y$；(2) 有些 X 是 Y 表示为 (a) $X \subset Y$ 或 (b) $Y \subset X$ 或 (c)$X \cap Y \neq \emptyset$；(3) 没有 X 是 Y 表示为 (d) $X \cap Y = \emptyset$；(4) 有些 X 不是 Y 表示为 (b) $Y \subset X$ 或 (c) $X \cap Y \neq \emptyset$ 或 (d) $X \cap Y = \emptyset$。

3 三段论式推理

三段论推理（Jeffrey，1981）是一种只有两个前提和三个项的演绎推理。一个非常著名的例子是所有希腊人都是人类。所有人类都是必死的。 $\therefore$ 所有希腊人都是必死的。 前提中共同的概念人类 确立了希腊人与必死的之间的关系。关系（或心理学文献中所称的“式”）分为：（1）全称肯定命题：所有 X 都是 Y；（2）特称肯定命题：有些 X 是 Y；（3）全称否定命题：没有 X 是 Y；（4）特称否定命题：有些 X 不是 Y。通过允许前提中的项互换位置，我们可以区分出256种不同类型的三段论推理（Khemlani 和 Johnson-

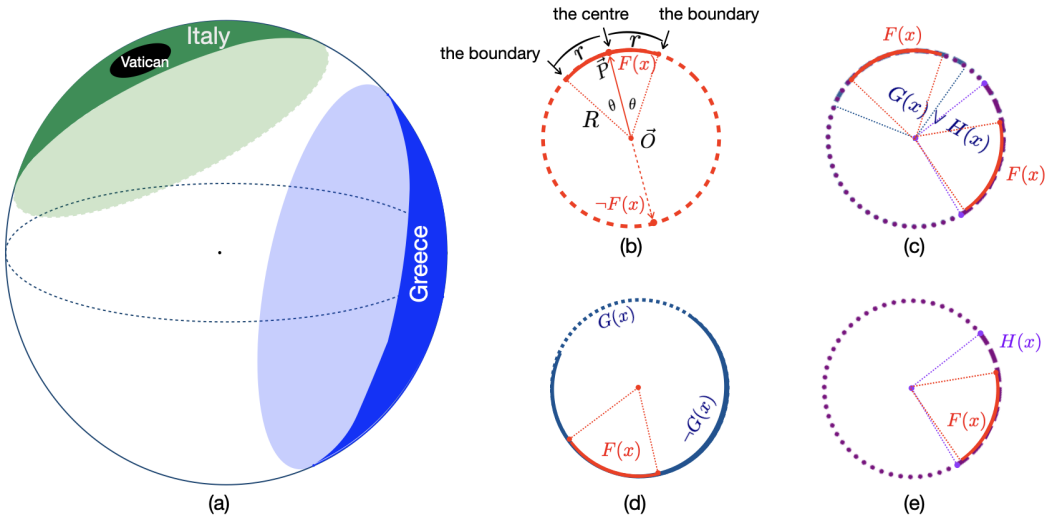

图4：（a）三维球体表面的欧拉图。（b）二维球体的表面是一个圆。该表面上的球体是一条曲线，例如，$F(x)$、$\neg F(x)$；（c）$\forall x \cdot F(x) \rightarrow G(x) \vee H(x)$：$F(x)$弧是$G(x)$弧或$H(x)$弧的一部分；（d）$\forall x \cdot F(x) \rightarrow \neg G(x)$：$F(x)$弧是$\neg G(x)$弧的一部分；（e）$\forall x \cdot F(x) \rightarrow H(x)$：$F(x)$弧是$H(x)$弧的一部分。

Laird，2012年）。这四种关系可以归纳为欧拉图形式的四种基本集合关系（Hammer 和 Shin，1998年）：（a）X 是 Y($X \subset Y$)的一部分，（b）X 包含 Y($Y \subset X$)，（c）X 与 Y($X \cap Y \neq \emptyset$)部分重叠，以及（d）X 与 Y($X \cap Y = \emptyset$)不相交，如图3所示。欧拉图可以在欧几里得空间（Dong等人，2024年）或双曲空间（Dong 等人，2025年）中构造成球体构型，并达到三段论推理的严谨性。该方法将集合 X 映射到球体 $\mathcal{O}_X$，并将四种三段论关系转换为以下四种空间关系。

- “所有 X 都是 Y” $\Leftrightarrow$ “球体 $\mathcal{O}_X$ 是球体 $\mathcal{O}_Y$的一部分”，**逆部分或包含**$(\mathcal{O}_X, \mathcal{O}_Y)$;
- “有些 X 是 Y” $\Leftrightarrow$ “球体 $\mathcal{O}_X$ 与球体 $\mathcal{O}_Y$不相互分隔”，$\neg\mathbf{D}(\mathcal{O}_X, \mathcal{O}_Y)$;
- “没有 X 是 Y” $\Leftrightarrow$ “球体 $\mathcal{O}_X$ 与球体 $\mathcal{O}_Y$相互分隔”，$\mathbf{D}(\mathcal{O}_X, \mathcal{O}_Y)$;
- “有些 X 不是 Y” $\Leftrightarrow$ “球体 $\mathcal{O}_X$ 不是球体 $\mathcal{O}_Y$的一部分”，$\neg$**逆部分或包含**$(\mathcal{O}_X, \mathcal{O}_Y)$.

在球面上高效地嵌入数据，可以解决旋转不变学习问题（Cohen 等，2018年；Esteves 等，2023年）。在这里，我们将概念嵌入为 n维球体表面上的圆，并将圆的构型构造成用于三段论推理的欧拉图。如果 $n = 3$，这些圆就是圆锥。举例来说，梵蒂冈位于意大利境内。意大利与希腊相互分隔。 $\therefore$ 梵蒂冈与希腊相互分隔。，其视觉示意如图4(a)所示。当我们在地球表面画一个圆时，从其中心到圆周的线沿着地球的一条大圆曲线延伸。形式化地，设 $\mathcal{O}$ 是一个以 $\vec{O}$ 为中心、半径为 R 的 n 维球体，$\vec{P}$ 是 $\mathcal{O}$ 表面上的一个点，$\|PO\| = R$。在 $\mathcal{O}$ 表面上，以 $\vec{P}$ 为中心、r 为半径的圆 $\bigcirc(\vec{P}, r)$ 定义为 $\mathcal{O}$ 表面上满足到 $\vec{P}$ 的曲面距离小于 r、$\arccos\left(\frac{\vec{P}}{\|\vec{P}\|} \cdot \frac{\vec{Q}}{\|\vec{Q}\|}\right) R < r$ 的点 $\vec{Q}$ 的集合。$\bigcirc(\vec{P}, r)$ 的补集也是一个圆，记作 $\bigcirc(\vec{P}', r')$，其中 $\vec{P}' = 2\vec{O} - \vec{P}$ 且 $r' = \left(\pi - \overline{\arccos}\left(\frac{\vec{P}}{\|\vec{P}\|} \cdot \frac{\vec{Q}}{\|\vec{Q}\|}\right)\right) R$。当 $n = 2$ 时，圆是弧，如图4(b)所示。圆的这一定义与（Dong等人，2024年，2025年）中的球体定义相容，可用于如下构建适用于一般三段论推理的欧拉图：设 a 是一个常量，x 是一个变量，F、G、H 是谓词。

- $F(a)$ 翻译为“原子圆 a 属于圆 F”，原子圆 a 具有最小半径 $\epsilon > 0$。形式上，**逆部分或包含**$(\bigcirc_a, \bigcirc_F)$，其中 $r_a = \epsilon$。
- $\forall x F(x) \rightarrow G(x)$ 可翻译为“对于任意原子圆 x，如果 x 是圆 F 的一部分，x 是圆 G 的一部分”。形式化地，我们写作 $\mathbf{P}(\bigcirc_F, \bigcirc_G)$。

- $\forall x F(x) \rightarrow G(x) \vee H(x)$。形式化地，我们写作 $\mathbf{P}(\bigcirc_F, \bigcirc_G) \vee \mathbf{P}(\bigcirc_F, \bigcirc_H)$。
- $\forall x F(x) \rightarrow \neg G(x)$。形式上，我们写作 **逆部分或包含**$(\bigcirc_F, \overline{\bigcirc_G})$。
- $\exists x F(x) \rightarrow G(x)$ 可译为“存在原子圆 a，若 a 是圆 F 的一部分，a 是圆 G 的一部分”。形式上，我们写作 **逆部分或包含**$(\bigcirc_a, \bigcirc_F) \wedge$ **逆部分或包含**$(\bigcirc_a, \bigcirc_G)$，其中 $r_a = \epsilon$。

We c在 n 维球体上，为以下 16 种三段论式推理类型创建圆配置
推理（Betz 等，2021）（详情见补充材料）。

1
$\forall x F(x) \rightarrow G(x)$。
$F(a)$ 。
$G(a)$ 。∴

2
$\forall x F(x) \rightarrow \neg G(x)$ 。
$F(a)$ 。
$\neg G(a)$ 。∴

3
$\forall x F(x) \rightarrow G(x)$ 。
$\neg G(a)$。
$\neg F(a)$。 ∴

4
$\forall x F(x) \rightarrow \neg G(x)$。
$G(a)$。
$\neg F(a)$。 ∴

5
$\forall x F(x) \rightarrow \neg G(x)$。
$\forall x G(x) \rightarrow \neg F(x)$。∴

6
$\forall x F(x) \rightarrow G(x)$。
$\forall x \neg G(x) \rightarrow \neg F(x)$。∴

7
$\forall x F(x) \rightarrow G(x)$。
$\forall x G(x) \rightarrow H(x)$。
$\forall x F(x) \rightarrow H(x)$。∴

8
$\forall x F(x) \rightarrow \neg G(x)$。
$\forall x \neg G(x) \rightarrow H(x)$。
$\forall x F(x) \rightarrow H(x)$。∴

9
$\forall x F(x) \rightarrow G(x)$。
$\forall x \neg H(x) \rightarrow \neg G(x)$。
$\forall x F(x) \rightarrow H(x)$。∴

10
$\forall x F(x) \rightarrow \neg G(x)$。
$\forall x \neg H(x) \rightarrow G(x)$。
$\forall x F(x) \rightarrow H(x)$。 ∴

11
$\forall x F(x) \rightarrow G(x)$。
$\exists x H(x) \rightarrow \neg G(x)$。
$\exists x H(x) \rightarrow \neg F(x)$。 ∴

12
$\forall x \neg F(x) \rightarrow G(x)$。
$\exists x H(x) \rightarrow \neg G(x)$。
$\exists x H(x) \rightarrow F(x)$。∴

13
$\forall x F(x) \rightarrow G(x) \vee H(x)$】【。
$\forall x F(x) \rightarrow \neg G(x)$。
$\forall x F(x) \rightarrow H(x)$。∴

14
$\forall x F(x) \rightarrow G(x) \vee H(x)$】【。
$\forall x G(x) \rightarrow \neg F(x)$。
$\forall x F(x) \rightarrow H(x)$。∴

15
$\forall x F(x) \rightarrow G(x) \vee H(x)$.
$\forall x G(x) \rightarrow J(x)$。
$\forall x H(x) \rightarrow J(x)$。
$\forall x F(x) \rightarrow J(x)$。 ∴

16
$\forall x F(x) \rightarrow G(x) \vee H(x)$】【。
$\forall x J(x) \rightarrow \neg G(x)$。
$\forall x J(x) \rightarrow \neg H(x)$。
$\forall x F(x) \rightarrow \neg J(x)$。 ∴

4 用于可靠析取三段论推理的球体神经网络 g

我们的新型球面神经网络（SphNN）在三个方面扩展了 SphNN（Dong等人，2024年、2025年）：(1) 我们将球体定义为 n 维球面上的圆；(2) 我们允许使用补集，并为补集定义补圆；(3) 我们的 SphNN 通过逐一显式构造三段论式陈述 f_i 的圆形配置，来确定析取范式 $f_1 \vee ... \vee f_n$ 中逻辑公式的可满足性，其中 f_i 为合取形式 $g_1 \wedge ... \wedge g_m$，$g_i$ 限定为以下形式之一：**逆部分或包含**$(\bigcirc_X, \bigcirc_Y)$，¬**逆部分或包含**$(\bigcirc_X, \bigcirc_Y)$，**D**$(\bigcirc_X, \bigcirc_Y)$，¬**D**$(\bigcirc_X, \bigcirc_Y)$，以及**逆部分或包含**$(\bigcirc_X, \bigcirc_Y)$。

确定三段论陈述可满足性的主要过程是一个逐步下降的过程，始于所有圆彼此重合。配置通过遵循邻域转换图迭代更新。在球面上构造圆形配置时，邻域转换图与（Dong等人，2024年、2025年）中的结构相同，但要求两个圆的直径之和小于球体大圆的周长，如图5(a)所示。这确保我们的 SphNN 继承了原始 SphNN 的确定性推理 特性，具体如下。

> 对于任何可满足的三段论陈述，*SphNN* 都能在一个训练周期内以全局损失为零的状态，正确构造球体构型作为欧拉图。

因此，如果 SphNN 在第一个训练周期后仍未能构造出目标图，就会得出结论：输入的三段论陈述不可满足，且结论的否定成立（算法1概述见补充材料）。为了判断罐B中是否有葡萄，我们的 SphNN 将反驳“罐B中没有葡萄”这一假设。具体做法是无法为该假设构造圆配置，即 $\mathbf{P}(\bigcirc_g, \bigcirc_A) \vee \mathbf{P}(\bigcirc_g, \bigcirc_B)$,

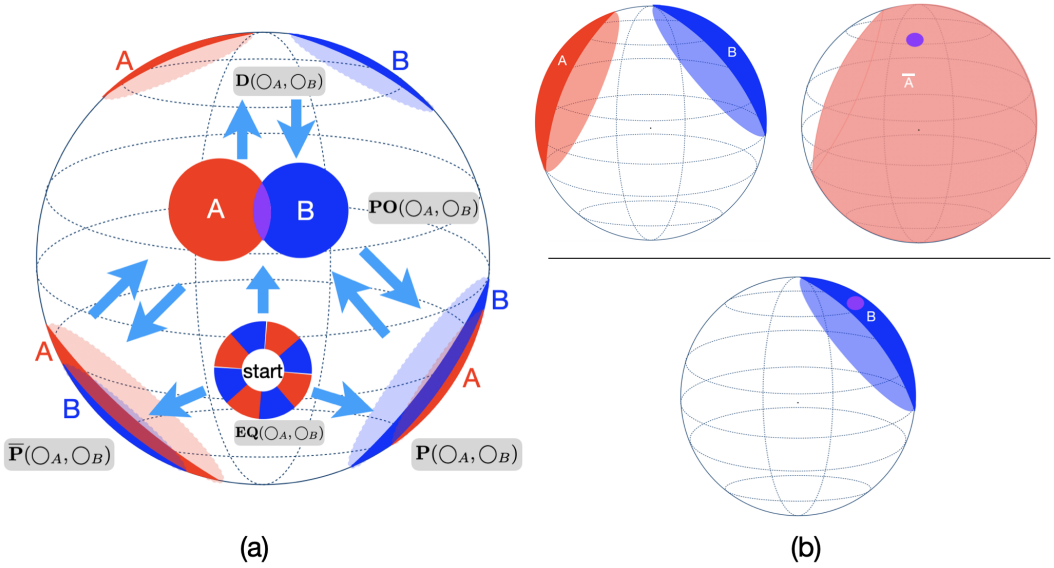

图5：(a) 球面上圆之间的邻域转换图。**PO**表示“部分重叠”，**EQ**表示“相等”，**P** 表示“逆部分或包含”；(b) 析取三段论推理的圆配置：（前提1）葡萄 $\bigcirc_g$ 要么在罐子A中，要么在罐B中，$\mathbf{P}(\bigcirc_g, \bigcirc_A) \vee \mathbf{P}(\bigcirc_g, \bigcirc_B)$；（前提2）葡萄不在罐子A中，$\mathbf{P}(\bigcirc_g, \bigcirc_A)$；（结论）葡萄在罐B中，$\mathbf{P}(\bigcirc_g, \bigcirc_B)$。

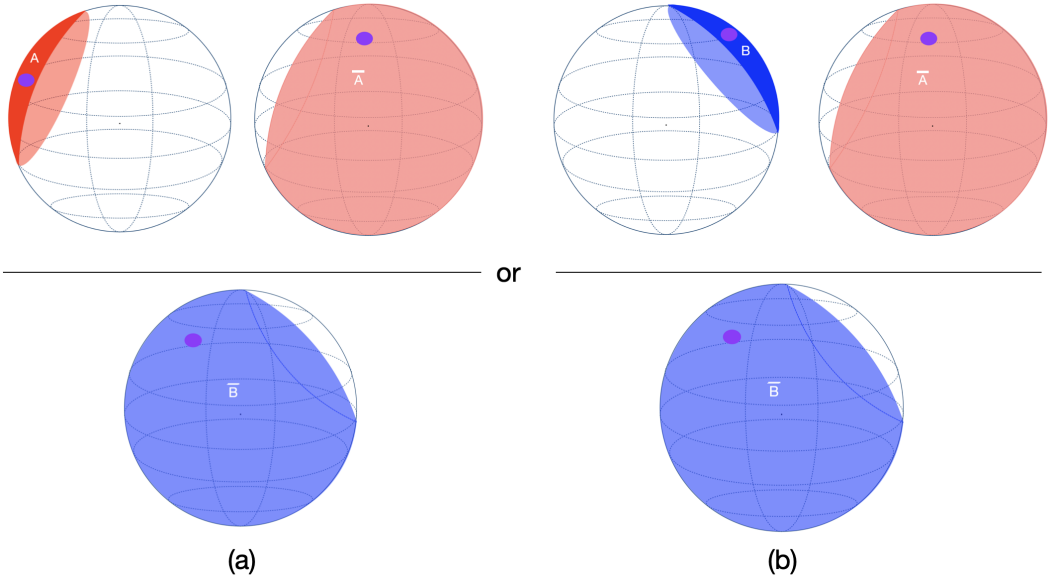

图6：(a) 情况1：葡萄在罐子A中；(b) 情况2：葡萄在罐B中。如果我们的 SphNN 无法为这两种情况都构造出配置，就会得出结论：葡萄肯定在罐B中。

$\mathbf{P}(\bigcirc_g, \overline{\bigcirc_A})$，以及 $\mathbf{P}(\bigcirc_g, \overline{\bigcirc_B})$。这等同于以下两种情况：(a) $\mathbf{P}(\bigcirc_g, \bigcirc_A)$，$\mathbf{P}(\bigcirc_g, \bigcirc_A)$，$\mathbf{P}(\bigcirc_g, \overline{\bigcirc_B})$；以及 (b) $\mathbf{P}(\bigcirc_g, \bigcirc_B)$，$\mathbf{P}(\bigcirc_g, \overline{\bigcirc_A})$，$\mathbf{P}(\bigcirc_g, \overline{\bigcirc_B})$，如图6所示。我们的 SphNN 会尝试为每种情况构造圆配置。在两种情况下都失败后，它就会确信葡萄在罐B中。

5 个实验

我们开展了四项实验，使用各种训练和测试数据集比较 SphNN 与欧拉网络。在实验 1 中，我们展示了 SphNN 在 16 种三段论式推理类型中实现了符号级严谨性（Betz 等，2021）。在实验2中，我们表明 SphNN 在经典三段论推理中仍能保持严谨性。在实验3中，我们针对析取三段论推理重新训练欧拉网络（Wang 等人，2018、2020），使其达到 100%准确率；在实验 4 中，我们表明，在经典三段论推理中达到 100.00% 的欧拉网络可以通过训练，在析取三段论推理中达到 100%准确率。然而，此后它在经典三段论推理中的表现会下降；当输入图像的模式与训练数据中的模式不同时，其表现还会进一步下降。

5.1 实验 1

数据集 我们将第3节中的16种三段论式推理转换为析取范式下的圆配置（详情见补充材料），并生成另外三种三段论结论，共计64种不同的圆形配置。在 $16 \times 4 = 64$种三段论推理陈述中，有32种有效。

方法 为了判断三段论推理是否有效，SphNN 会尝试在球面上构造一个反例。若构造成功，SphNN 则判定该三段论推理无效；否则，SphNN 判定原结论有效。

实验设置 圆的初始半径为 e^{-1}。三个半径为 e^{-1} 的圆被随机初始化为重合状态，位于半径为 1 的球面上。我们将学习率设为 0.0001，最大训练轮数为 $M = 1$。所有实验均在一台配备 Apple M1 Max（10 核 CPU/24 核 GPU）和 32 GB 内存的 MacBook Pro 上进行。我们让 SphNN 尝试构造以下维度的庞加莱球：2，3，15，30，100，200，1000，2000，3000，10000。

结果 SphNN通过构造维度从2到10000的圆配置，成功判定了32种有效和32种无效的三段论推理类型，共计640项推理任务。判定有效推理的平均耗时为108.04秒。为无效推理构造反例的平均耗时为7.07秒；320个无效推理案例中有221个（69.06%）在5秒内完成判定；320个有效案例中有226个（70.06%）在120秒内完成判定。

5.2 实验2

数据集 我们使用SphNN的数据集（Dong等人，2025年）。

方法和实验设置 与实验 1相同。

结果 我们的SphNN通过构造维度从2到10000的圆配置，成功判定了24种有效和232种无效的三段论推理类型，共计2560项推理任务。判定有效推理的平均耗时为64.00秒。为无效推理构造反例的平均耗时为11.27秒；2320个无效推理案例中有2137个（92.11%）在5秒内完成判定；240个有效案例中有199个（82.92%）在120秒内完成判定。

5.3 与监督式深度学习的比较研究

开发神经网络三段论推理极具挑战性，曾一度被视为乌托邦式的设想（Khemlani和Johnson-Laird，2012年）。直到最近，研究人员才开发出监督式神经网络欧拉网络，以近似实现相当大一部分三段论推理（Wang 等人，2018、2020），如图 7(a)所示。EN 的输入是两幅图像，每幅图像由两个带有集合论关系的彩色圆组成。圆的颜色用于区分三段论推理中的三个项。两幅输入图像中相同的颜色代表中项。欧拉网络通过两个孪生网络，将每幅输入图像编码为一个潜在向量。EN 的输出是一个表示

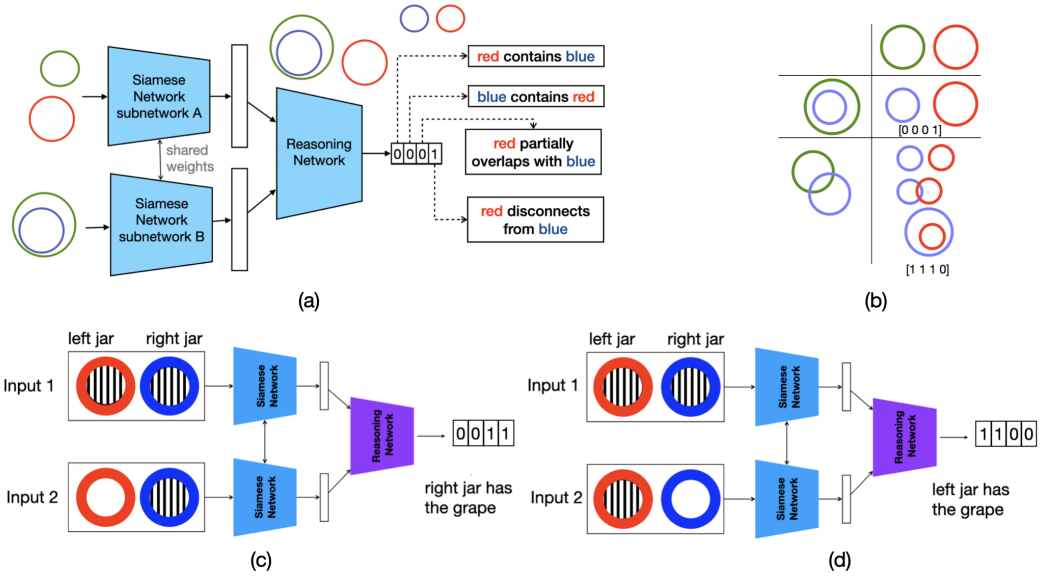

图 7：（a）欧拉网络的架构；（b）欧拉网络的组合表。如果两个前提是“蓝色在绿色内部。绿色与红色分离”，则组合结果仅为“蓝色与红色分离”，表示为 $[0,0,0,1]$。（c、d）欧拉网络用于析取三段论推理的训练数据。

主项与谓项之间集合论关系的向量。将两个前提映射为结论的方式列于组合表中，可能的结论以向量表示，如图 7(b)所示。训练数据的形式为 ((图像, 图像), 向量)。欧拉网络可以经过训练，以处理经典三段论推理的各种变体（Wang 等人，2018年）。

5.3.1 实验3

数据集 我们将欧拉网络（Wang 等人，2018）改用于析取三段论推理。训练数据由两个前提和一个结论向量组成，用来表示“葡萄在罐子里”的情境。第一张输入图像包含两个覆盖着条纹的圆，以表示不确定性。圆的颜色、粗细和相对位置保持固定，而圆心位置则随机变化。第二张输入图像与第一张相同，只是随机选取其中一个圆设为空，表示该圆不含葡萄。输出向量编码析取推理的结论：如果葡萄在左侧圆中，则为 $[1100]$；如果葡萄在右侧圆中，则为 $[0011]$，如图 7（c、d）所示。

实验设置 按照（Wang 等人，2018）的实验设置，我们创建了 96000 条输入输出记录，其中 88000 条记录用于训练，8000 条记录用于验证和测试。

结果 欧拉网络在我们的测试数据集上进行析取三段论推理时达到了 100%准确率。

5.3.2 实验 4

我们考察输入模式的变化是否会影响欧拉网络的推理表现，以及欧拉网络能否在经典三段论推理中继续取得较高表现。

数据集 我们为析取三段论推理设计了三种变体输入模式，如图 8(a-c)所示。第一种变体中，圆的颜色是随机化的；第二种变体中，颜色和粗细都经过随机化；第三种变体中，第一个输入保持不变，但第二个输入显示一个空圆和一个填满的圆。我们生成了经典三段论推理的测试数据。所有实验的测试数据集均与实验3相同。

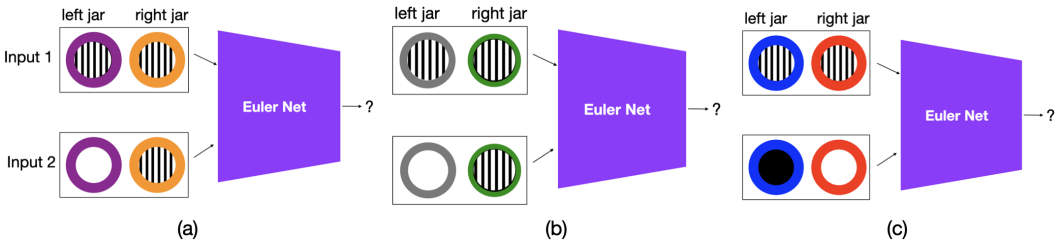

图 8：三个变体测试数据集对实验3中训练良好的欧拉网络构成挑战 .

结果 随着视觉特征变化，准确率有所下降：颜色随机化时为75.00%，颜色和粗细均变化时为53.57%，采用替代的清晰/完整输入模式时为46.43%。对于经典三段论推理，其表现下降至6.25%。

6 结论与展望

析取三段论推理是一种基础的决策策略，使人类、动物和人工智能系统能够通过排除错误候选项得出结论。本文比较了三种用于基础决策的神经网络，展示了通过显式模型构建进行推理这一方法的优势。

1. 这种方法能够构建可由人类解释的模型。相比之下，监督式深度学习系统和大语言模型都是黑箱。2. 这种方法不需要训练数据，因为前提包含得出结论所需的全部信息。在实践中，这可以节省大量资源，包括数据、训练时间和能源。从理论上说，这种方法完全消除了分布外问题。3. 这种方法支持持续学习——新学会的构建方法可以与旧方法共存。相比之下，监督式深度学习系统会遭受严重的灾难性遗忘。欧拉网络学会析取三段论推理后，将无法执行已经学会的经典三段论推理。4. 这种方法利用邻域转换图构建模型——为某一任务设计的图，只要具备描述目标构型的表示能力，就可以用于其他任务。我们展示了，为析取三段论推理设计的图无需任何微调，即可直接用于另外15种三段论式推理。相比之下，监督式深度学习系统需要针对每项任务使用特定的训练数据。它们或许可以通过微调进行迁移学习。5. 这种方法实现了符号级三段论式逻辑推理，为解决人工智能文献中符号主义与连接主义之间的冲突开辟了一条新途径。

球体是表示部分—整体关系的直观视觉工具。这种关系在日常生活中无处不在，也是各门科学中分类系统的基本结构原则（Smith，1996）。许多关系概念都可以归结为区域之间的连接关系，这一传统在哲学和心理学领域由来已久（de Laguna，1922；Whitehead，1929；Piaget，1954；Smith，1994；Dong，2008）。例如，“A 比 B 高”可以定义如下：设 X 是地平面上方的一个区域。

1. A 站在地面上时可以与 X 相连；2. B 站在地面上
时不能与 X 相连。

方位可以归结为距离比较（Dong 和 Guesgen，2008）。“左边”的基本含义源于与心脏的接近程度。“B 在 C 的左边”可以解释如下：如果你将身体朝向 C 投影，B 会更靠近你身体中心脏所在的一侧。我们可以引入模糊球体边界，使 SphNN 能够表示渐变隶属度和模糊概念（Zadeh，1965）。

球体是通过引入一个非零半径对向量的推广。反过来，传统神经网络可以视为一种特殊的球面神经网络，其中所有球体半径都设为零。未来的研究方向包括探索 SphNN 与监督式深度学习和大语言模型的无缝集成，将其应用于认知建模，以及为高风险领域开发可解释且安全的人工智能模型。

参考文献

Abramson, J., Adler, J., Dunger, J. 等（2024年）。利用 AlphaFold 3 准确预测生物分子相互作用的结构。《自然》。Ambags, E. L., Capitoli, G., Imperio, V. L., Provenzano, M., Nobile, M. S. 和 Li`o, P.（2023年）。利用模糊概率决策树辅助临床实践。*arXiv*。Betz, G., Voigt, C. 和 Richardson, K.（2021年）。语言模型的批判性思维。载于第十四届国际计算语义学会议（*IWCS*）论文集，第63–75页，荷兰格罗宁根（线上）。计算语言学协会。Biever, C.（2023年）。人工智能聊天机器人的简易智能测试。《自然》，619：686–689。Cohen, T. S., Geiger, M., K¨ohler, J. 和 Welling, M.（2018年）。球面卷积神经网络。载于国际学习表征会议。Constant, A.（2023年）。法律三段论推理的贝叶斯模型。《人工智能与法律》，32：1–22。Creswell, A., Shanahan, M. 和 Higgins, I.（2022年）。选择—推断：利用大语言模型实现可解释的逻辑推理。de Laguna, T.（1922年）。点、线与面作为实体集合。《哲学杂志》，19：449–461。Deng, W., Pei, J., Kong, K., Chen, Z., Wei, F., Li, Y., Ren, Z., Chen, Z. 和 Ren, P.（2023年）。用于法律判决分析的三段论推理。载于 Bouamor, H., Pino, J. 和 Bali, K.（编），*2023*年自然语言处理实证方法会议论文集，新加坡。计算语言学协会。Dong, T.（2008年）。关于 RCC 的评论：从 RCC 到 RCC^{++}。《哲学逻辑杂志》，37(4)：319–352。Dong, T. 和 Guesgen, H.（2008年）。基于距离比较的方位关系统一框架。载于第七届 *IEEE* 认知信息学国际会议论文集，第75–82页。IEEE计算机学会出版社，斯坦福大学，美国加利福尼亚州。Dong, T., Jamnik, M. 和 Li`o, P.（2024年）。用于理性推理的球体神经网络。*arXiv:2403.15297*。Dong, T., Jamnik, M. 和 Li`o, P.（2025年）。通过构建可解释模型实现可靠的神经推理。载于*AAAI*。Eisape, T., Tessler, M., Dasgupta, I., Sha, F., van Steenkiste, S. 和 Linzen, T.（2024年）。对人类与语言模型的三段论推理进行系统比较。载于北美计算语言学协会年会。Esteves, C., Slotine, J.-J. 和 Makadia, A.（2023年）。扩展球面卷积神经网络。载于 Krause, A., Brunskill, E., Cho, K., Engelhardt, B., Sabato, S. 和 Scarlett, J.（编），第*40*届国际机器学习会议论文集，第202卷，机器学习研究论文集，第9396–9411页。PMLR。Fedorenko, E., Piantadosi, S. T. 和 Gibson, E. A. F.（2024年）。语言主要是沟通工具，而非思维工具。《自然》，630：575–586。Ferrigno, S., Huang, Y. 和 Cantlon, J. F.（2021年）。猴子运用析取三段论进行推理。《心理科学》，32(2)：1–9。

Goodwin, G. 和 Johnson-Laird, P.（2005年）。关系推理。《心理学评论》，112：468–93。Google（2023年）。PaLM 2 技术报告。Hammer, E. 和 Shin, S. J.（1998年）。欧拉视觉逻辑。《逻辑史与逻辑哲学》，19(1)。He, K., Zhang, X., Ren, S. 和 Sun, J.（2016年）。用于图像识别的深度残差学习。第770–778页。Jeffrey, R.（1981年）。《形式逻辑：范围与局限（第二版）》。纽约州纽约市：McGraw-Hill。Jiang, A. Q., Sablayrolles, A., Mensch, A., Bamford, C., Chaplot, D. S., de las Casas, D., Bressand, F., Lengyel, G., Lample, G., Saulnier, L., Lavaud, L. R., Lachaux, M.-A., Stock, P., Scao, T. L., Lavril, T., Wang, T., Lacroix, T. 和 Sayed, W. E.（2023年）。Mistral 7B。Johnson-Laird, P. N. 和 Byrne, R. M. J.（1991年）。演绎。Lawrence Erlbaum Associates, Inc. Khemlani, S. 和 Johnson-Laird, P. N.（2012年）。三段论理论：荟萃分析。《心理学公报》，138(3)：427–457。Knauff, M.（2009年）。基于心智模型与视觉意象的演绎关系推理神经认知理论。《空间认知与计算》，9(2)：109–137。Knauff, M.（2013年）。《推理的空间》。MIT出版社。Knauff, M., Fangmeier, T., Ruff, C. C. 和 Johnson-Laird, P. N.（2003年）。推理、模型与意象：行为测量与皮层活动。《认知神经科学杂志》，15(4)：559–573。Lampinen, A. K., Dasgupta, I., Chan, S. C. Y., Sheahan, H. R., Creswell, A., Kumaran, D., McClelland, J. L. 和 Hill, F.（2024年）。语言模型与人类一样，在推理任务中表现出内容效应。《*PNAS Nexus*》，3(7)。Mistral, A. T.（2023年）。专家混合模型：高质量稀疏专家混合模型。Mitchell, M.（2023年）。我们如何知道人工智能系统有多聪明？《科学》，381(6654)：adj5957。Mushlin, S. B. 和 Harry L. Greene, I.（2010年）。《医学决策——算法方法》。爱思唯尔，美国费城。

OpenAI（2023年）。GPT-3.5。

Park, P. S., Goldstein, S., O'Gara, A., Chen, M., 和 Hendrycks, D.（2024年）。人工智能欺骗：案例、风险与潜在解决方案综述。*Patterns*，5(5)。Piaget, J.（1954年）。儿童现实的建构。Routledge & Kegan Paul Ltd. Raschka, S.（2024年）。从零开始构建大语言模型。Manning。Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lock- hart, E., Hassabis, D., Graepel, T., Lillicrap, T., 和 Silver, D.（2020年）。通过基于学习模型进行规划，精通雅达利、围棋、国际象棋和将棋。《自然》，588:604—609。Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., 和 Hassabis, D.（2017年）。在没有人类知识的情况下精通围棋。《自然》，550:354–359。Smith, B.（1994年）。认知科学的拓扑学基础。载于 Eschenbach, C., Habel, C., 和 Smith, B. 编，认知科学的拓扑学基础。纽约州布法罗。FISI-CS 研讨会。Smith, B.（1996年）。部分整体拓扑学：部分与边界理论。数据与知识工程，20:287–303。

Soliman, A.（2024年）。Deepmind 的人工智能天气预报系统击败世界一流系统。《自然》，636:282–283。

Touvron, H., Martin, L., Stone, K., Albert, P., Almahairi, A., Babaei, Y., Bashlykov, N., Batra, S., Bhargava, P., Bhosale, S., Bikel, D., Blecher, L., Ferrer, C. C., Chen, M., Cucurull, G., Esiobu, D., Fernandes, J., Fu, J., Fu, W., Fuller, B., Gao, C., Goswami, V., Goyal, N., Hartshorn, A., Hosseini, S., Hou, R., Inan, H., Kardas, M., Kerkez, V., Khabsa, M., Kloumann, I., Korenev, A., Koura, P. S., Lachaux, M.-A., Lavril, T., Lee, J., Liskovich, D., Lu, Y., Mao, Y., Martinet, X., Mihaylov, T., Mishra, P., Molybog, I., Nie, Y., Poulton, A., Reizenstein, J., Rungta, R., Saladi, K., Schelten, A., Silva, R., Smith, E. M., Subramanian, R., Tan, X. E., Tang, B., Taylor, R., Williams, A., Kuan, J. X., Xu, P., Yan, Z., Zarov, I., Zhang, Y., Fan, A., Kambadur, M., Narang, S., Rodriguez, A., Stojnic, R., Edunov, S., and Scialom, T.（2023年）。Llama 2：开放基础模型与微调聊天模型。Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., and Polosukhin, I.（2017）。注意力就是你所需要的一切。载于 Guyon, I., Luxburg, U. V., Bengio, S., Wal- lach, H., Fergus, R., Vishwanathan, S., and Garnett, R. 编，神经信息处理系统进展，第 30 卷。Curran Associates, Inc. Wang, D., Jamnik, M., and Li`o, P.（2018年）。利用深度神经网络研究图示推理。载于 *Diagrams 2018*，第 390–398 页。Wang, D., Jamnik, M., and Li`o, P.（2020年）。利用多路复用图网络进行抽象图示推理。载于国际学习表征会议。Whitehead, A. N.（1929年）。《过程与实在》。Macmillan Publishing Co., Inc. Wysocka, M., Carvalho, D., Wysocki, O., Valentino, M., and Freitas, A.（2025年）。SylloBio-NLI：评估大语言模型的生物医学三段论推理能力。*arXiv:2410.14399*。Yamins, D. L. and DiCarlo, J. J.（2016）。使用目标驱动的深度学习模型理解感觉皮层。《自然·神经科学》，19(3):356–365。Zadeh, L. A.（1965）。模糊集。《信息与控制》，8:338–353。Zelikman, E., Wu, Y., Mu, J., and Goodman, N.（2022年）。STar：通过推理进行自举。在 Oh, A. H., Agarwal, A., Belgrave, D., and Cho, K. 编，神经信息处理系统进展。

A 使用 GPT-5 系列进行的补充实验

在第 1 节中，我们综述了近期对大语言模型在三段论推理（包括析取三段论）中的推理表现所做的评估。尽管这些研究一致认为，大语言模型仍无法完美地进行三段论推理，但我们仍可乐观地期待它们在不久的将来能够做到这一点。因此，我们开展了补充实验，旨在找出一种能够在经典三段论推理中达到 100%准确率的大语言模型及与之匹配的输入形式。我们设计了以下四种输入模式，共计 $256 \times 4 = 1024$ 项经典三段论推理任务。

> 1. 使用有意义词语的三段论陈述，例如：所有希腊人都是人类；2. 概念由两个词语组成、并以“ ”连接的三段论陈述，例如：有些希腊人 山地人不是人类 番茄；3. 使用简单符号的三段论陈述，例如：有些 S 是 P；4. 使用长随机符号的三段论陈述，例如：有些 *X2asfi3f23* 是 *aH21¨u01qH*。

我们让 OpenAI GPT-5 和 GPT-5-nano 判断三段论推理是否可满足，并在可能的情况下构造欧拉图。

结果是，GPT-5 只有在输入长随机符号时，判断可满足性才能达到 100%准确率。然而，在这些正确决策中，GPT-5 约有 10% 给出了错误的

欧拉图。从 GPT-5-nano 到 GPT-5，在减少“决策正确但解释错误”的情况方面并没有显著改进。神经网络的训练过程通常会在达到 100%准确率时停止，这表明 GPT-5 的决策极不可靠。详情见表1。

表1：OpenAI GPT-5-nano 和 GPT-5 的三段论推理表现。‘expl’表示解释正确，‘wrg’表示解释错误。‘#correct decision-wrg’列表示决策正确但解释错误；‘#wrong decision-expl’列表示决策错误但解释正确。

版本	表层形式	#正确决策		#错误决策		#简单准确率
		expl	wrg	expl	wrg	
GPT-5-nano	词语	160（62.5%）	90	5	1	97.7%
	双词	230（89.4%）	22	4	0	98.4%
	简单符号	226（88.3%）	24	6	0	97.7%
	长随机符号	222（86.7%）	25	9	0	96.5%
gpt-5	词语	239 (93.4%)	16	1	0	99.6%
	双词	234（91.4%）	21	1	0	99.6%
	简单符号	236（92.2%）	15	5	0	98.0%
	长随机符号	231（90.2%）	25	0	0	**100.0%**

B 16种推理类型的圆形配置

我们在（Betz等人，2021年）中将 n维球体表面的圆形配置列为16种三段论式推理类型的欧拉图，具体如下。

1. 广义假言推理

$$\begin{array}{l}\forall x F(x) \rightarrow G(x). \\ F(a). \\ \hline G(a). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 是圆 G 的一部分。原子圆 a 是圆 F 的一部分。原子圆 a是圆 G 的一部分。”，如下所示。

all F are G, all a are F, all a are G

另外三个三段论结论如下。

all F are G, all a are F, no a are G

all F are G, all a are F, some a are G

all F are G, all a are F, some a are not G

2. 广义假言推理的否定变体

$$\begin{array}{l}\forall x F(x) \rightarrow \neg G(x). \\ F(a). \\ \hline \neg G(a). \quad \therefore\end{array}$$

其圆配置表述为：“圆 F 是圆 G 的补圆的一部分。原子圆 a 是圆 F 的一部分。原子圆 a 是圆 G 的补圆的一部分。”如下所示。

all F are c_G, all a are F, all a are c_G

我们用“c_”表示“.....的补圆”。另外三个三段论结论如下。

all F are c_G, all a are F, no a are c_G

all F are c_G, all a are F, some a are c_G

all F are c_G, all a are F, some a are not c_G

3. 广义换位

$$\frac{\forall x F(x) \to \neg G(x).}{\forall x G(x) \to \neg F(x). \quad \therefore}$$

其圆配置可表述为“圆 F 是圆 G 的补圆的一部分。圆 G 是圆 G 的补圆的一部分。”，如下所示。

all F are c_G, all G are c_F

另外三个三段论结论如下。

all F are c_G, no G are c_F

all F are c_G, some G are c_F

all F are c_G, some G are not c_F

4. 广义换位的否定变体

$$\frac{\forall x F(x) \to G(x).}{\forall x \neg G(x) \to \neg F(x). \quad \therefore}$$

其圆配置可表述为“圆 F 是圆 G 的一部分。圆 F 的补圆是圆 F 的补圆的一部分。”，如下所示。

all F are G, all c_G are c_F

另外三个三段论结论如下。

all F are G, no c_G are c_F

all F are G, some c_G are c_F

all F are G, some c_G are not c_F

5. 假言三段论1

$$\frac{\begin{array}{l}\forall x F(x) \to G(x). \\ \forall x G(x) \to H(x).\end{array}}{\forall x F(x) \to H(x). \quad \therefore}$$

其圆配置可读作：“圆 F 是圆 G 的一部分。圆 G 是圆 H 的一部分。圆 F 是圆 H 的一部分”，如下所示。

all F are G, all G are H, all F are H

另外三个三段论结论如下。

all F are G, all G are H, no F are H

all F are G, all G are H, some F are H

all F are G, all G are H, some F are not H

6. 假言三段论1的否定变体

$$\frac{\begin{array}{l}\forall x F(x) \to \neg G(x). \\ \forall x \neg G(x) \to H(x).\end{array}}{\forall x F(x) \to H(x). \quad \therefore}$$

其圆形配置读作：“圆 F 是圆 G 的一部分。圆 G 是圆 H 的一部分。圆 F 是圆 H 的一部分”，如下所示。

all F are c_G, all c_G are H, all F are H

另外三个三段论结论如下。

all F are c_G, all c_G are H, no F are H

all F are c_G, all c_G are H, some F are H

all F are c_G, all c_G are H, some F are not H

7. 假言三段论 2

$$\begin{array}{l}\forall x F(x) \rightarrow G(x). \\ \forall x \neg H(x) \rightarrow \neg G(x). \\ \hline \forall x F(x) \rightarrow H(x). \quad \therefore\end{array}$$

其圆配置表述为：“圆 F 是圆 G 的一部分。圆 H 的补圆是圆 G 的补圆的一部分。圆 F 是圆 H 的一部分”，如下所示。

all F are G, all c_H are c_G, all F are H

另外三个三段论结论如下。

all F are G, all c_H are c_G, no F are H

all F are G, all c_H are c_G, some F are H

all F are G, all c_H are c_G, some F are not H

8. 假言三段论 2 的否定变体

$$\begin{array}{l}\forall x F(x) \rightarrow \neg G(x). \\ \forall x \neg H(x) \rightarrow G(x). \\ \hline \forall x F(x) \rightarrow H(x). \quad \therefore\end{array}$$

其圆配置表述为：“圆 F 是圆 G 的补圆的一部分。圆 H 的补圆是圆 G 的一部分。圆 F 是圆 H 的一部分”，如下所示。

all F are c_G, all c_H are G, all F are H

另外三个三段论结论如下。

all F are c_G, all c_H are G, no F are H

all F are c_G, all c_H are G, some F are H

all F are c_G, all c_H are G, some F are not H

9. 假言三段论 3

$$\begin{array}{l}\forall x F(x) \rightarrow G(x). \\ \exists x H(x) \rightarrow \neg G(x). \\ \hline \exists x H(x) \rightarrow \neg F(x). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 是圆 G 的一部分。原子圆 a 是圆 H 的一部分。原子圆 a 是圆 G 的补圆的一部分。原子圆 a 是圆 F 的补圆的一部分”，如下所示。

all F are G, all a are H, all a are c_G, all a are c_F

另外三个三段论结论如下。

all F are G, all a are H, all a are c_G, no a are c_F

all F are G, all a are H, all a are c_G, some a are c_F

all F are G, all a are H, all a are c_G, some a are not c_F

10. 假言三段论 3 的否定变体

$$\begin{array}{l}\forall x\neg F(x) \to G(x).\\ \exists x H(x) \to \neg G(x).\\ \hline \exists x H(x) \to F(x). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 的补圆是圆 G 的一部分。原子圆 a 是圆 H 的一部分。原子圆 a 是圆 G 的补圆的一部分。原子圆 a是圆 F 的一部分”，如下所示。

all c_F are G, all a are H, all a are c_G, all a are F

另外三个三段论结论如下。

all c_F are G, all a are H, all a are c_G, no a are F

all c_F are G, all a are H, all a are c_G, some a are F

all c_F are G, all a are H, all a are c_G, some a are not F

11. 广义否定后件式

$$\begin{array}{l}\forall x F(x) \to G(x).\\ \neg G(a).\\ \hline \neg F(a). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 是圆 G 的一部分。原子圆 a 是圆 F 的补圆的一部分。原子圆 a 是圆 F 的补圆的一部分。”，如下所示。

all F are G, all a are c_G, all a are c_F

另外三个三段论结论如下。

all F are G, all a are c_G, no a are c_F

all F are G, all a are c_G, some a are c_F

all F are G, all a are c_G, some a are not c_F

12. 广义否定后件式的否定变体

$$\begin{array}{l}\forall x F(x) \to \neg G(x).\\ G(a).\\ \hline \neg F(a). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 是圆 G 的补圆的一部分。原子圆 a 是圆 F 的一部分。原子圆 a 是圆 F 的补圆的一部分。”写作如下。

all F are c_G, all a are G, all a are c_F

其他三个三段论结论如下。

all F are c_G, all a are G, no a are c_F

all F are c_G, all a are G, some a are c_F

all F are c_G, all a are G, some a are not c_F

13. 析取三段论

$$\begin{array}{l}\forall x F(x) \to G(x) \vee H(x). \\ \forall x F(x) \to \neg G(x). \\ \hline \forall x F(x) \to H(x). \quad \therefore\end{array}$$

其圆配置可表述为："圆 F 属于圆 G 或圆 H。圆 F 属于圆 G 的补圆。圆 F 属于圆 H。"写法如下。

all F are G_or_H, all F are c_G, all F are H

另外三个三段论结论如下。

all F are G_or_H, all F are c_G, no F are H

all F are G_or_H, all F are c_G, some F are H

all F are G_or_H, all F are c_G, some F are not H

14. 析取三段论的否定变体

$$\begin{array}{l}\forall x F(x) \to G(x) \vee H(x). \\ \forall x G(x) \to \neg F(x). \\ \hline \forall x F(x) \to H(x). \quad \therefore\end{array}$$

其圆配置可表述为："圆 F 属于圆 G 或圆 H。圆 G 属于圆 F 的补圆。圆 F 属于圆 H。"写法如下。

all F are G_or_H, all G are c_F, all F are H

另外三个三段论结论如下。

all F are G_or_H, all G are c_F, no F are H

all F are G_or_H, all G are c_F, some F are H

all F are G_or_H, all G are c_F, some F are not H

15. 广义两难推理

$$\begin{array}{l}\forall x F(x) \to G(x) \vee H(x). \\ \forall x G(x) \to J(x). \\ \forall x H(x) \to J(x). \\ \hline \forall x F(x) \to J(x). \quad \therefore\end{array}$$

其圆形配置可读作："圆 F 是圆 G 或圆 H 的一部分。圆 G 是圆 J 的一部分。圆 H 是圆 J 的一部分。圆 F 是圆 J 的一部分。"写作如下。

all F are G_or_H, all G are J, all H are J, all F are J

另外三个三段论结论如下。

all F are G_or_H, all G are J, all H are J, no F are J

all F are G_or_H, all G are J, all H are J, some F are J

all F are G_or_H, all G are J, all H are J, some F are not J

16. 广义两难推理的否定变体

$$\begin{array}{l}\forall x F(x) \to G(x) \vee H(x). \\ \forall x J(x) \to \neg G(x). \\ \forall x J(x) \to \neg H(x). \\ \hline \forall x F(x) \to \neg J(x). \quad \therefore\end{array}$$

其圆配置可表述为：“圆 F 是圆 G 或圆 H 的一部分。圆 G 是圆 G 的补圆的一部分。圆 J 是圆 J 的补圆的一部分。圆 F 是圆 J 的补圆的一部分。”写法如下。

所有 F 都是 G 或 H，所有 J 都是 c G，所有 J 都是 c H，所有 F 都是 c J

另外三个三段论结论如下。

所有 F 都是 G 或 H，所有 J 都是 c G，所有 J 都是 c H，没有 F 是 c J 所有 F 都是 G 或 H，所有 J 都是 c G，所有 J 都是 c H，某些 F 是 c J 所有 F 都是 G 或 H，所有 J 都是 c G，所有 J 都是 c H，某些 F 不是 c J

C算法

给定目标圆配置，我们将其转换为析取范式 $f_1 \vee \cdots \vee f_m$。每个f_i 都是圆关系的合取形式。对于每个 f_i，我们搜索一个圆的环链（在链中，可以出现一个圆或它的补圆，但不能两者同时出现），并让 SphNN 判断这个圆环配置是否可满足。如果一个 f_i 可满足，我们的 SphNN 就会得出结论：$f_1 \vee \cdots \vee f_m$ 可满足。例如，配置 $\mathbf{P}(\bigcirc_A, \bigcirc_B)$、$\mathbf{P}(\bigcirc_B, \bigcirc_C)$ 和 $\neg\mathbf{D}(\bigcirc_A, \overline{\bigcirc_C})$，表示圆 A、圆 B 的补集和圆 C 之间的一种关系。

算法1：SphNN 的主过程 输入：n维球体上的目标圆配置。输出：输入配置可满足或不可满足 1 将输入配置转换为析取范式 $f_1 \vee \cdots \vee f_m$；2 Loss $\leftarrow +\infty$；3 foreach $f_i \in [f_1 \ldots f_m]$ do 4 圆环 获取一组来自 f_i的圆；5 if 圆环 $== \emptyset$then 6 continue 7 else 8 Loss $\leftarrow$ 为圆环构建一个配置；9 if Loss $== 0$ then 10 break 11 if Loss $== 0$then 12 return 可满足 13 else 14 return 不可满足

D 代码和数据

代码和数据可按要求提供。